SINGLE-MOLECULE STUDY OF THYMIDINE GLYCOL AND I-MOTIF THROUGH THE ALPHA-HEMOLYSIN ION CHANNEL

by

Lidong He

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry

The University of Utah

December 2013

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Lidong He**

has been approved by the following supervisory committee members:

Cynthia J. Burrows	, Chair	**08/12/2013**
		Date Approved
Henry S. White	, Member	**08/12/2013**
		Date Approved
Matthew S. Sigman	, Member	**08/12/2013**
		Date Approved

and by **Cynthia J. Burrows** , Chair/Dean of

the Department/College/School of **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Nanopore-based devices have emerged as a single-molecule detection and analysis tool for a wide range of applications. Through electrophoretically driving DNA molecules across a nanosized pore, a lot of information can be received, including unfolding kinetics and DNA-protein interactions. This single-molecule method has the potential to sequence kilobase length DNA polymers without amplification or labeling, approaching “the third generation” genome sequencing for around $1000 within 24 hours. α-Hemolysin biological nanopores have the advantages of excellent stability, low-noise level, and precise site-directed mutagenesis for engineering this protein nanopore.

The first work presented in this thesis established the current signal of the thymidine glycol lesion in DNA oligomers through an immobilization experiment. The thymidine glycol enantiomers were differentiated from each other by different current blockage levels. Also, the effect of bulky hydrophobic adducts to the current blockage was investigated.

Secondly, the α-hemolysin nanopore was used to study the human telomere i-motif and RET oncogene i-motif at a single-molecule level. In Chapter 3, it was demonstrated that the α-hemolysin nanopore can differentiate an i-motif form and single-strand DNA form at different pH values based on the same sequence. In

addition, it shows potential to differentiate the folding topologies generated from the same DNA sequence.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

α-HL	α-hemolysin
Btn	biotin
CD	circular dichroism
DMT	dimethoxytrityl
DPhPC	1,2-diphytanoyl-*sn*-glycero-3-phospho-choline
ddH_2O	double-distilled water
dsDNA	double-stranded DNA
Gh	guanidinohydantoin
GNM	glass nanopore membrane
HPLC	high performance liquid chromatography
mC	5-methylcytosine
MS	mass spectrometry
MspA	*Mycobacterium smegmatis* porin A
ODN	oligodeoxynucleotide
ROS	reactive oxygen species
ssDNA	single-stranded DNA
Sp	spiroiminodihydantoin
TBA	thrombin-binding aptamer

TBDMS	tert-butyldimethylsilyl
TCAA	trichloroacetic acid
Tg	thymidine glycol
UV-Vis	ultra-violet visible spectroscopy

ACKNOWLEDGEMENTS

I would like to express my great gratitude to my research advisor, Prof. Cynthia J. Burrows, who taught me what research is all about and inspired me during my research. The 2 years of research in her lab greatly increased my knowledge in the biological chemistry field. Her enthusiasm and dedication to science, and her understanding and support have made me a better researcher for my future career.

I have great admiration for my graduate committee, Prof. Matthew S. Sigman, Prof. Janis Louie, Prof. Thomas E. Cheatham, and most especially, Prof. Henry S. White, who also mentored me in the collaboration with his patience and intelligence. Special thanks to Dr. Aaron Fleming, who provided numerous insightful ideas during our discussion and shared invaluable research experience with me. I would like to thank Dr. James Muller for his support and advice in my research.

The Burrows group members and White group members have been very helpful, and my research life would be much less fun without them: Dr. Jan Riedl, Dr. Na An, Dr. Qian Jin, Dr. John Watkins, Dr. Robert Johnson, Dr. Uday Ghanty, Dr. Khiem Nguyen, Pranjali Ghude, Ania Wolna, Yun Ding, Xibo Li, Omar Alshykhly, Long Luo, Rukshan Perera, Shuo Tan, Anton Alenko, and Srinivas Paladugu.

Last but not the least, I am grateful to my parents. They have provided me with their unconditional love, understanding, and great support on my education.

CHAPTER 1

INTRODUCTION

Single-molecule techniques possess many advantages over ensemble studies, as one molecule can be observed at a time to study individual dynamics and distribution instead of averaging signals (*1*). Emerging as a promising single-molecule method, biological ion channels have been under intensive investigation for their use in the study of biomolecules (*2*). α-Hemolysin (α-HL) ion channel is excreted from the pathogen *Staphylococcus aureus* as a toxic protein channel. This toxin can form a transmembrane ion channel to release vital molecules from the host cell such as heme (*3*). As shown in Figure 1.1, α-HL is a mushroom-shaped heptameric complex that can be self-assembled in a phospholipid bilayer. It contains two parts: a vestibule that has a *cis*-entrance opening with a diameter of ~2.6 nm, a cavity with a diameter of ~4.5 nm, and a β-barrel that has a *trans*-exit with a 2.2 nm diameter. The narrowest part of this α-HL is referred to as the constriction zone with a diameter around 1.4 nm (*3*). The White laboratory developed a fabrication method to produce a glass nanopore membrane (GNM) to support lipid bilayers (*4*). and α-HL can self-assemble across this lipid bilayer, forming the ion channel for the current recording system (Figure 1.2). Until now, a wide range of molecules have been studied within α-HL, including nucleic acids,

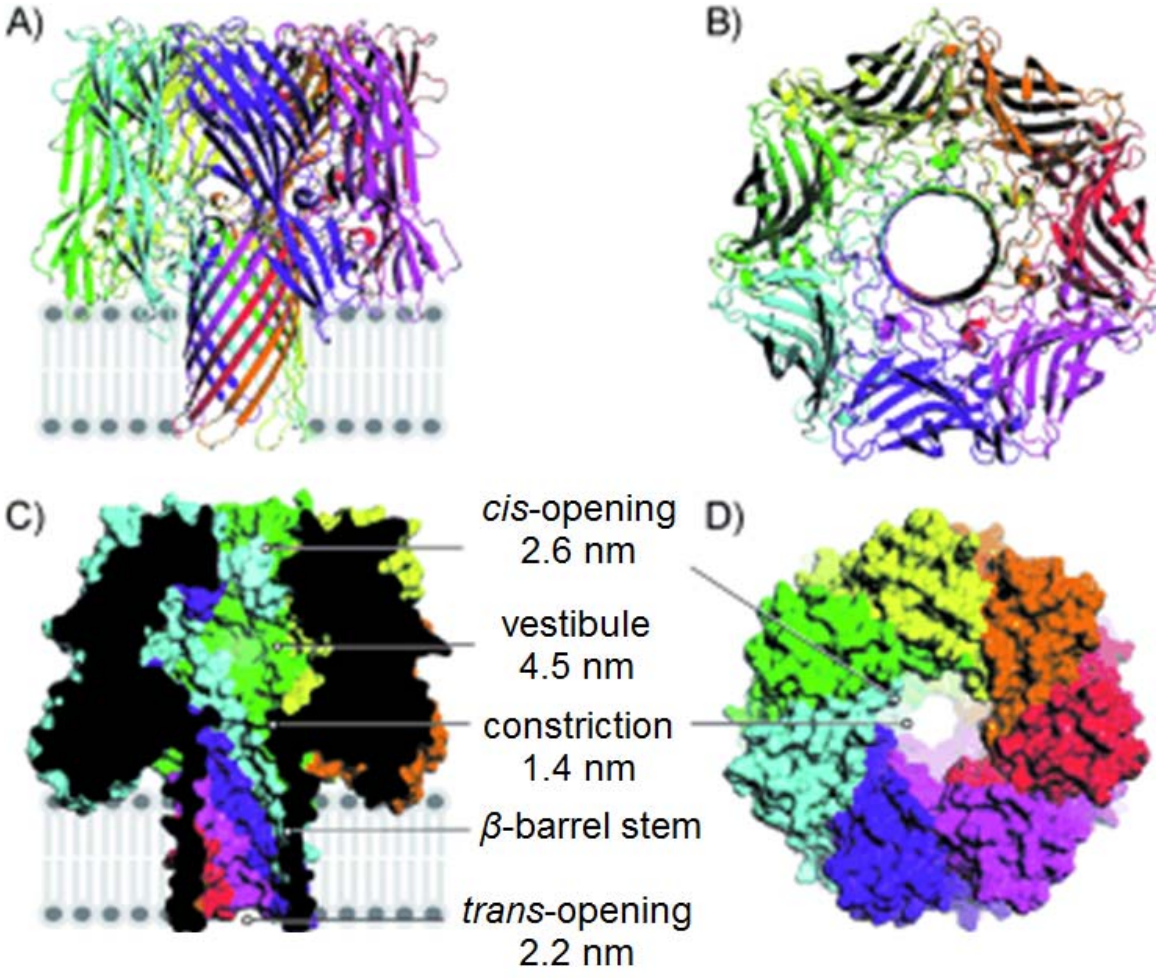


Figure 1.1. α-HL heptameric complex side view (A), top view (B), and dimensions of this ion channel (C and D). Figure adapted from Ma *et al.* (*5*).

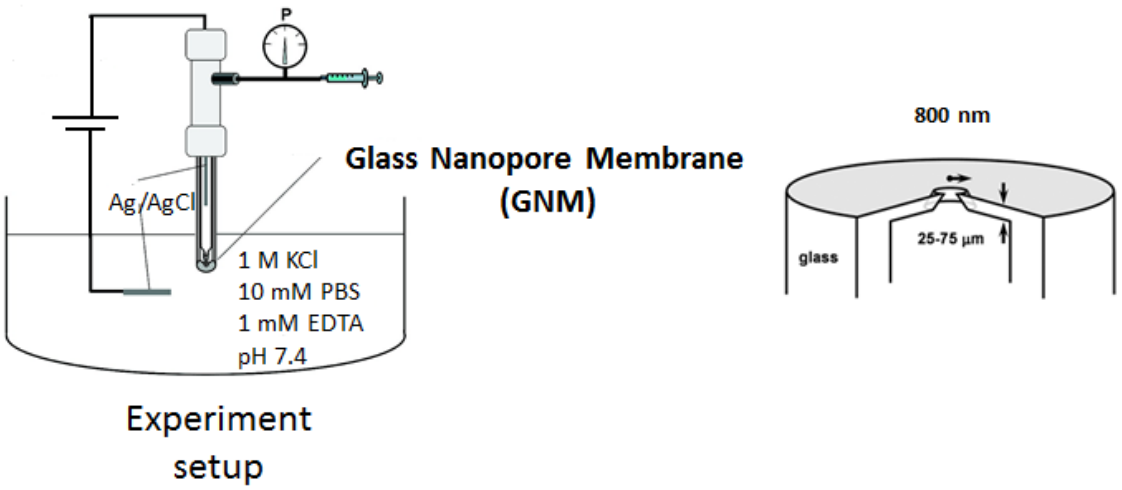


Figure 1.2. Schematic for the nanopore experiment set-up that was used to conducte ion channel current recordings. A GNM was used to support a lipid bilayer, across which the α-HL ion channel can form. Two Ag/AgCl electrodes were placed inside the GNM capillary and in the cell buffer solution, respectively.

small molecules and metal-protein interactions (*5*–*7*).

1.1 α-HL Nanopore for DNA Sequencing

As shown in Figure 1.1, the size of the constriction zone (1.4 nm) in the stem allows only single-stranded DNA (ssDNA) to pass through because it has a diameter of ~1 nm, while double-stranded DNA (dsDNA) has a diameter of ~2 nm (*8*). Once driven by electrophoretic force, the negatively charged DNA translocates through the α-HL nanopore, causing a decrease in ion flux through the ion channel. The duration of the translocation of ssDNA corresponds to the length of the DNA, while the current blockage is determined by many factors such as size, hydrophobicity, shape, and charge of the DNA (*9*).

α-HL nanopore is proposed as a promising tool for the $1000 Genome Project (http://www.genome.gov/12513210). Ideally, four different DNA native bases (adenine, thymine, guanine, and cytosine) can be distinguished based on their current-time trace, which provides sequencing information for the translocated DNA strand (*10*). However, the fast translocation time (~1–20 μs per nucleotide) and the small current level differences between the four native bases (~1–2 pA, depending on the applied bias) make the current technique not sensitive enough to differentiate the four bases (*10*).

Different approaches have been proposed and tested to address the above challenges. By covalently attaching cyclodextrin within the α-HL nanopore β-barrel, individual nucleotides can be captured and translocated through this engineered α-HL nanopore, during which it binds to cyclodextrin, generating four specific ion

current blockages corresponding to the four nucleotides, respectively (*11*). Therefore, it has been suggested that the α-HL cap region can be adapted with an exonuclease for use in DNA sequencing (Figure 1.3) (*10*). Once this α-HL attached exonuclease binds dsDNA, it will cleave nucleotides one at a time, releasing the cleaved nucleotides into the α-HL nanopore followed by interaction with the cyclodextrin adaptor, generating the corresponding current signal. However, for sequencing purposes, it is important that 100% exonuclease-cleaved nucleotides be captured by cyclodextrin and in the same order as the exonuclease cleaved the strand, which is a challenge for this method (*10*).

A second approach aims at decreasing the DNA translocation speed through the protein nanopore. As shown in Figure 1.3, phi29 DNA polymerase was used to control DNA templates through forward and reverse ratcheting. This method decreased DNA translocation rates to 2.5–40 nucleotides per second and with one nucleotide precision (*12*). However, the main issue for this method is that specific types of DNA damage, such as thymidine glycol, can stall the DNA polymerase, resulting in termination of the DNA sequencing process.

Before the alpha-hemolysin ion channel nanopore method is successfully used as a DNA sequencing tool, the current signature for each nucleotide must be obtained. Currently, the best method for obtaining nucleotide current signatures is through an immobilization experiment (*13*). This method allows the DNA to stay in the pore for a much longer time, and therefore, the current level of a substituted base at a specific position in the sensing region can be measured (Figure 1.4). The current blockage of all four native DNA bases was achieved through immobilizing

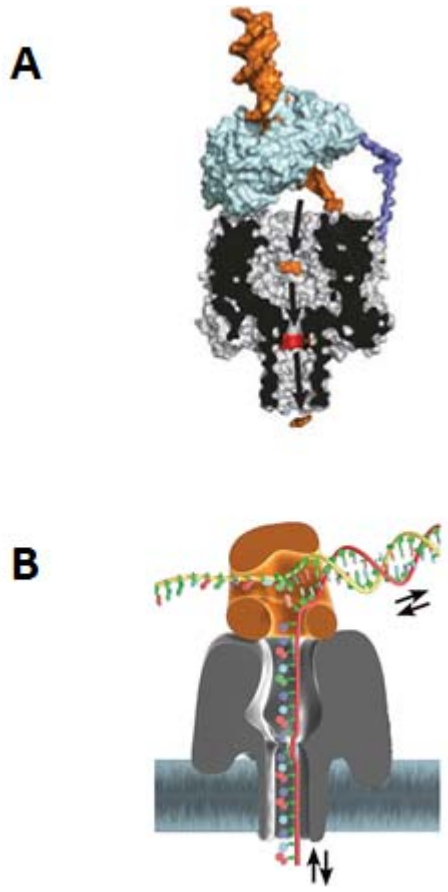


Figure 1.3. A cartoon shows an exonuclease attached to α-HL and cleaved nucleotides captured by cyclodextrin (A). DNA is ratcheted forward and reversely through the α-HL nanopore (B). Figure adapted from Branton *et al.* and Schneider *et al*. (*10, 14*).

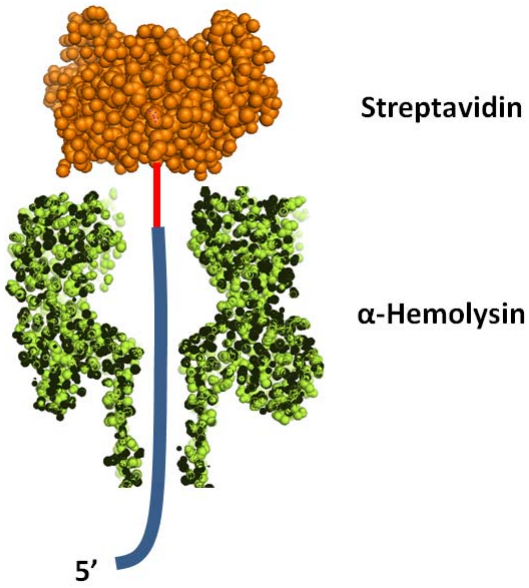


Figure 1.4. Representation of streptavidin (PDB 1MK5), biotinylated DNA and α-HL (PDB 7AHL) for the immobilization experiment. α-HL nanopore sensing zone in the β-barrel region.

a streptavidin-biotin ssDNA in the nanopore (*13*). Besides the native bases, DNA is susceptible to damage by a wide range of exogenous and endogenous agents, generating many forms of base damage. Therefore, for DNA sequencing, the current signal of these DNA damaged nucleotides needs to be established to avoid errors during sequence reading.

In Chapter 2, an immobilization experiment was applied to establish the current signal of thymidine glycol. Thymidine glycol (Tg) is an important oxidative damage biomarker (*15*), therefore, it is very significant to study the current blockage level caused by Tg within the α-HL ion channel. Through immobilization experiments, it was found that ω_{15} (ω represents numbering from the nontraditional 3'-end) position is the most sensitive current blockage site for Tg. The current blockage of (5*R*, 6*S*)-Tg and (5*R*, 6*R*)-Tg was differentiated in the current blockage histogram. This established Tg current signal will facilitate construction of a nucleotide current signal database used for DNA sequencing.

1.2 α-HL Nanopore for DNA Secondary Structures Study

α-HL nanopore has been utilized to study DNA hairpins. Akeson's group reported that single-base pair resolution was achieved by plotting residual current I/I_0 versus duration to differentiate 3 to 8 base pair hairpins (*16*). Additionally, an α-HL nanopore was used to study 9 base pair blunt-ended hairpins (*16*). By analysis of ionic current signals of hairpins with different terminal base pairs, the Akeson group concluded that the different current blockage levels correspond to different orientations of the hairpin in the α-HL vestibule. A working model has been

proposed to explain four different current signatures produced by the blunt ended hairpins (*17*).

α-HL nanopore has also been utilized to investigate G-Quadruplexes which are folded tertiary structures formed from G-rich sequences. α-HL nanopore was used to study the folding and unfolding kinetics of thrombin-binding aptamer (TBA) G-quadruplex, and cation selective G-quadruplex stability at the single-molecule level (*18*). Also, it was recently shown that human telomere G-quadruplex hybrid1 and hybrid 2 folds are well-differentiated through their interaction with α-HL ion channel in our laboratory (*19*).

In Chapter 3, α-HL ion channel was demonstrated to detect i-motif structures which are tertiary structures formed from C-rich sequences. The current blockage signal exhibited differences between a single-stranded DNA form and an i-motif form under different pH values based on the same sequence. The current blockage histogram can be utilized to analyze different i-motif conformations. Compared to ensemble experiments, α-HL ion channel has the potential to differentiate topologies of the i-motif. This information will facilitate design of drugs that target the i-motif, since nanopore analysis can provide the most populated i-motif conformation in solution for targeting by small molecules and potentially differentiate the bound and unbound forms of the i-motif complexes.

1.3 References

1. Kapanidis, A. N., and Strick, T. (2009) Biology, one molecule at a time, *Trends Biochem. Sci. 34*, 234–243.

2. Reiner, J. E., Balijepalli, A., Robertson, J. W. F., Campbell, J., Suehle, J., and Kasianowicz, J. J. (2012) Disease detection and management via single nanopore-based sensors, *Chem. Rev. 112*, 6431–6451.

3. Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., Gouaux, J. E. (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore, *Science. 274*, 1859–1866.

4. Zhang, B., Galusha, J., Shiozawa, P. G., Wang, G., Bergren, A. J., Jones, R. M., White, R. J., Ervin, E. N., Cauley, C. C., and White, H. S. (2007) Bench-top method for fabricating glass-sealed nanodisk electrodes, glass nanopore electrodes, and glass nanopre membranes of controlled size, *Anal. Chem. 79*, 4778–4787.

5. Ma, L., and Cockroft, S. L. (2009) Biological nanopores for single-molecule biophysics, *ChemBioChem 11*, 25–34.

6. Kang, X. F., Cheley, S., Guan, X., and Bayley, H. (2006) Stochastic detection of enantiomers, *J. Am. Chem. Soc. 128*, 10684–10685.

7. Choi, L., Mach, T., and Bayley, H. (2013) Rates and stoichiometries of metal ion probes of cysteine residues within ion channels, *Biophys. J. 105*, 356–364.

8. Deamer, D. W., Branton, D. (2002) Characterization of nucleic acids by nanopore analysis, *Acc. Chem. Res. 35*, 817–825.

9. Wolna, A., Fleming, A. M., An, N., He, L., White, H. S., and Burrows, C. J. (2013) Electrical current signatures of DNA base modifications in single molecules immobilized in the α-hemolysin ion channel, *Isr. J. Chem 53*, 417–430.

10. Branton, D., Deamer, D. W., Marziali, A., Bayley, H., Benner, S. A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A., Huang, X., Jovanovich, S. B., Krstic, P., S., Lindsay, S., Ling, X. S., Mastrangelo, C. H., Meller, A., Oliver, J. S., Pershin, Y. V., Ramsey, J. M., Riehn, R., Soni, G. V., Tabard-Cossa, V., Wanunu, M., Wiggin, M., and Schloss, J. A. (2008) The potential and challenges of nanopore sequencing, *Nat. Biotechnol. 26*, 1146–1153.

11. Clarke, J., Wu, H., Jayasinghe, L., Patel, A., Reid, S., and Bayley, H. (2009) Continuous base identification for single-molecule nanopore DNA sequencing, *Nat. Nanotechnol. 4*, 265–270.

12. Cherf, G. M., Lieberman, K. R., Rashid, H., Lam, C. E., Karplus, K., and Akeson, M. (2012) Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision, *Nat. Biotechnol. 30*, 344–348.

13. Stoddart, D., Heron, A. J., Mikhailova, E., Maglia, G., and Bayley, H. (2009) Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore, *Proc. Natl. Acad. Sci. U.S.A. 106*, 7702–7707.

14. Scheneider, G. F. and Dekker, C. (2012) DNA sequencinng with nanopores, *Nat. Biotechnol. 30*, 326–328.

15. Adelman, R., Saul, R. L., Ames, B. N. (1988) Oxidative damage to DNA: relation to species metabolic rate and life span, *Proc. Natl. Acad. Sci. U.S.A. 85*, 2706–2708.

16. Vercoutere, W., Winters-Hilt, S., Olsen, H., Deamer, D., Haussler, D., and Akeson, M. (2001) Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel, *Nat. Biotechnol. 19*, 248–252.

17. DeGuzman, V. S., Lee, C. C., Deamer, D. W., and Vercoutere, W. A. (2006) Sequence-dependent gating of an ion channel by DNA hairpin molecules, *Nucleic Acids Res. 34*, 6425–6437.

18. Shim, J. W., Tan, Q., and Gu, L. (2009) Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity, *Nucleic Acids Res. 37*, 972–982.

19. An, N., Fleming, A. M., and Burrows, C. J. (2013) Interactions of the human telomere sequence with the nanocavity of the α-hemolysin ion channel reveal structure-dependent electrical signatures for hybrid folds, *J. Am. Chem. Soc. 135*, 8562–8570.

CHAPTER 2

SINGLE-MOLECULE DETECTION OF THYMIDINE GLYCOL THROUGH THE ALPHA-HEMOLYSIN ION

2.1 Introduction

Oxidation is one of the most common forms of damage to human DNA. It is mainly induced by reactive oxygen species (ROS), which are released during cellular respiration, and the processes of biosynthesis and biodegradation (*1*). Thymidine glycol (5,6-dihydro-5,6-dihydroxythymidine, Tg) is the most common oxidized thymidine lesion found after exposure with ROS generated by ionizing radiation (*2*), oxidizing agents such as potassium permanganate (*3*), or as a consequence of aerobic metabolism (*4*). As shown in Figure 2.1, thymidine glycol contains two chiral centers, forming a mixture of four stereoisomers. It has been reported that γ-irradiation yields equal amounts of the 5*R* and 5*S* isomers, but the 5*R* isomer is more stable. For the 5*R* pair, the equilibrium ratio between the *cis* and *trans* forms is about 87% to 13% (*5*). It has been estimated that about 400 thymidine glycols are created per human cell per day (*4*). Thymidine glycol exerts significant distortion on the duplex DNA helix because the thymidine glycol base appears to be extrahelical, as shown by proton NMR spectroscopy (*6*). As a

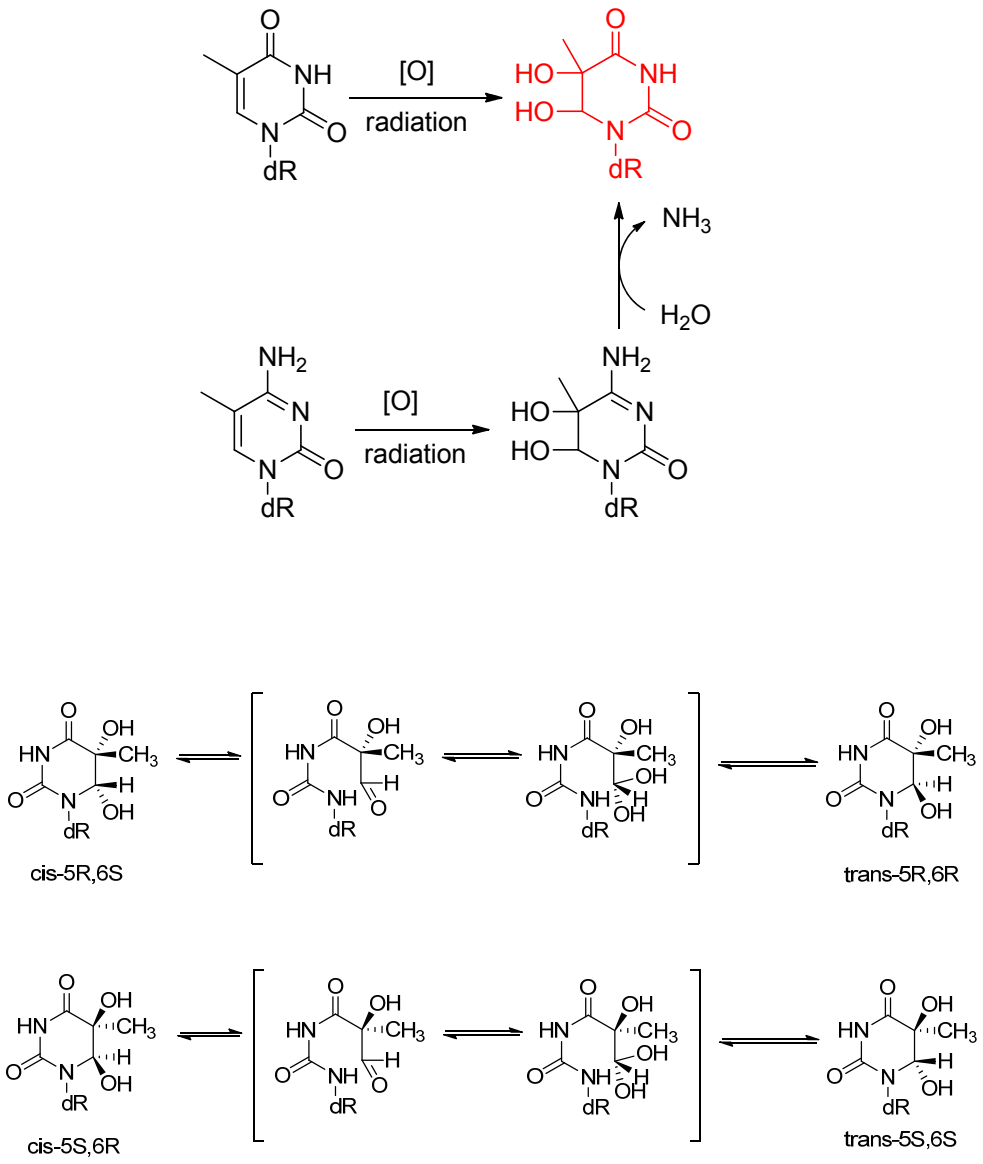


Figure 2.1. Thymidine glycol formation pathway and epimerization between the *cis* and *trans* forms of the 5*R* and 5*S* thymidine glycol diastereomers.

result, primer extension studies with DNA polymerase I with a template containing thymidine glycol showed that extension stops at the lesion site with 2'-deoxyadenine inserted opposite the thymidine glycol (*7*). Therefore, thymidine glycol effectively blocks DNA replication. Removal of thymidine glycol is accomplished by the base excision repair (BER) pathway, initiated by endonuclease III and MutM from *E. coli* (*8, 9*). If not repaired, it can prevent DNA replication and thus prevent cell proliferation (*10*). Therefore, there is a need to develop a rapid detection method for Tg damage in DNA.

Several DNA damage detection methods commonly used today include the comet assay (*11*) and HPLC-MS spectrometry (*12*). The comet assay provides high sensitivity and can detect the location of damage, but it cannot provide information on damage identity or multiple damage sites. HPLC-MS can distinguish the two isomers of thymidine glycol, but it cannot provide information on the location of the specific damaged site, and it is not cost efficient. All of the current thymidine glycol detection methods have the drawbacks of not being able to do single-molecule or multiple-lesion detection, and they cannot provide sequencing information in a cost-efficient manner.

Before the alpha-hemolysin (α-HL) ion channel nanopore method is successfully used as a DNA sequencing tool, the current signature for each individual nucleotide must be obtained. Thymidine glycol is an important oxidative damage biomarker; therefore, it is very significant to study the current blockage level caused by Tg within the α-HL ion channel. The best method for obtaining nucleotide current signatures is an immobilization experiment. In this experiment,

biotin is attached on the 3'-end of oligodeoxynucleotides (ODNs), and then streptavidin is added which helps ODNs suspend within the α-HL ion channel for a long time to record the current blockage level. In this work, based on nanopore immobilization experiments, the current blockage of Tg in an ODN was investigated and compared to the four native bases. ODNs containing Tg in positions ω13–16 (from 3' end) were studied because these positions were reported as the most sensitive sensing area within the wild-type α-HL pore (*13*). To investigate the effect of tandem Tg lesions to current blockage level, ODNs that contain double Tg damages ($Tg_{\omega 13,15}$, $Tg_{\omega 13,16}$) were also studied based on tethering experiments. In order to gain insight into how the interaction between DNA and the nanopore influences the current blockage, the TBDMS (tert-butyldimethylsilyl) protecting group was left on Tg in order to study the influence of the bulky hydrophobic group on the ion current level.

2.2 Experimental Section

2.2.1 DNA Preparation and Purification Procedures

(*5R, 6S*)-Tg phosphoramidites and native DNA phosphoramidites were purchased from Glen Research, Sterling, VA. The 40mer 3'-biotinylated ODNs with one Tg at positions ω13–16, respectively, were synthesized by the DNA-Peptide Core Facility at the University of Utah. Each ODN was cleaved from the synthetic column and deprotected according to the manufacturer's protocols after synthesis. Then the ODNs were purified by an ion-exchange HPLC column. The HPLC method was a linear gradient of 25% to 100% B over 25 minutes while monitoring

absorbance at 260 nm (A = 25 mM Tris, 1 M NaCl in 10% CH_3CN/90% ddH_2O, pH 7, B = 10% CH_3CN/90% ddH_2O, flow rate = 1 mL/min). After HPLC purification, the ODNs were dialyzed against ddH_2O to remove salts and then characterized by UV-Vis.

2.2.2 Formation of Streptavidin-Biotin ssDNA Complexes

3'-Biotinylated ODNs were mixed with streptavidin at a 4:1 ratio in a solution of 1 M KCl, 10 mM PBS, 1 mM EDTA (pH 7.4). Before ion channel recordings, streptavidin and 3'-biotinylated ODNs were incubated for 10 minutes at room temperature.

2.2.3 Chemicals and Materials for Nanopore Measurement

Buffer electrolyte was prepared with 1 M KCl, 10 mM PBS, and 1 mM EDTA (pH 7.4) solution. The buffer solution was then filtered with a sterile 0.22 μm Millipore vacuum filter. The phospholipid, 1,2-diphytanoyl-*sn*-glycero-3-phospho-choline (DPhPC), was purchased in a powder form from Avanti Polar Lipids and stored at -20 °C. Before conducting the nanopore experiment, DPhPC was dissolved in decane at 10 mg/mL for use. Wild-type α-hemolysin was dissolved in water at 1 mg/mL and stored at -80 °C. Glass nanopore membranes (GNMs) were fabricated as previously described and used as the support structure for the lipid bilayer (*14*). Before use, GNMs were chemically modified with a 2% (v/v) (3-cyano-propyl) dimethylchlorosilane in acetonitrile to produce a hydrophobic surface.

2.2.4 Current-Time Recordings and Data Analysis

A custom-built high-impedance, low-noise amplifier and data acquisition system (Electronic Bio Sciences, San Diego, CA) was used for the current-time recordings. Two Ag/AgCl electrodes were positioned inside and outside of the GNM capillary to apply a voltage. A pressure gauge and a 10 mL gas tight syringe (Hamilton) were attached to the GNM. First the DPhPC solution was painted on the GNM surface, followed by applying a positive pressure to the GNM capillary facilitating reconstitution of a single α-HL ion channel in the lipid bilayer. A good α-HL ion channel was identified by a resistance change from 50–100 GΩ (bilayer resistance) to ~1 GΩ (protein resistance).

In the immobilization studies, after protein reconstitution into the lipid bilayer, the ODN (after incubation with streptavidin) was added to the *cis* compartment of the nanopore cell (final concentration 5 μM). +120 mV was applied (*trans* vs. *cis*), and the streptavidin-biotin ssDNA complex was captured by the α-HL pore and caused an ion current change from an open channel current (I_0) to a blocking current level (I). Since streptavidin is too bulky to enter the pore, the blocking current was maintained for 750 ms before -120 mV was applied to eject the ODN from the nanopore for 200 ms. Then the voltage was returned back to +120 mV again to attract the next ODN (Figure 2.2). To compare the blocking current level of different ODNs, a C_{40} mer was added next as an internal standard. A minimum of 500 events were collected for each sample. The current-time traces were filtered at 10 kHz and sampled at 50 kHz.

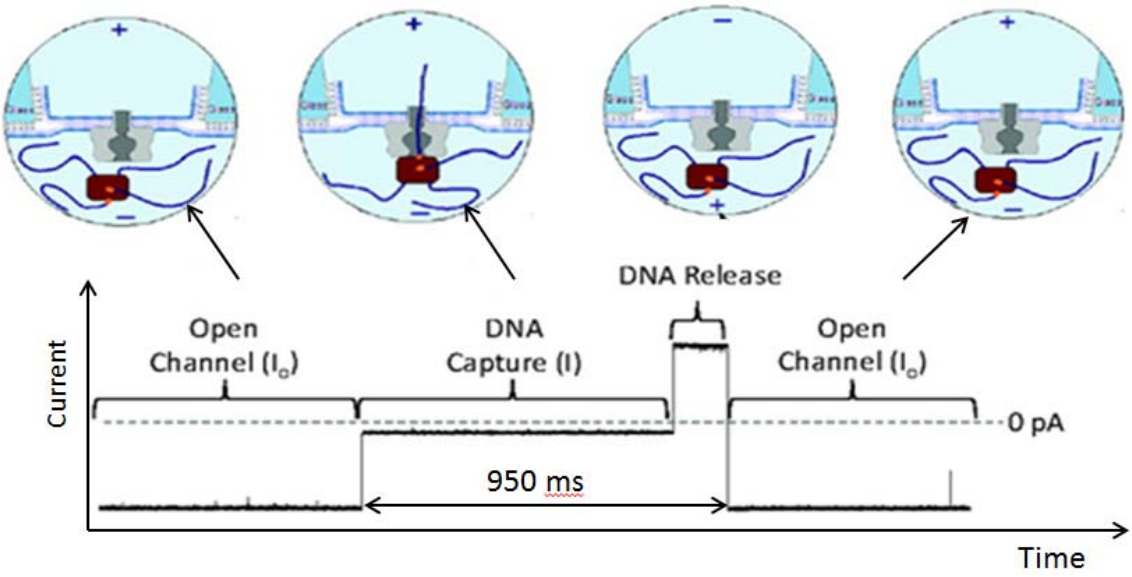


Figure 2.2. Representative current vs. time trace from the immobilization experiment. Figure adapted from Schibel *et al*. (*15*).

Events were extracted using QuB (version 1.5.0.31). Histograms of the current blockade were plotted using OriginPro (version 8.5.1). Density plots were analyzed with data analysis programs provided by EBS.

2.3 Results and Discussion

2.3.1 Ion Channel Recordings on the Tg-Containing ODNs

In order to understand how Tg influences the current blockage level in the α-HL ion channel, tethering experiments were conducted. A representative current–time trace from the tethering experiment is shown in Figure 2.2. To be consistent with previous results in our laboratory, a 3'-biotinylated C_{40} strand was used in this experiment as an internal standard, and its I/I_0 was set to 0. 3'- Biotinylated native DNA $C_{36}T_4$ was used as a control for the following experiment. The $\Delta\%I/I_0$ current histogram shows $C_{36}T_4$ is 1.1% less blocking than C_{40} (Figure 2.3).

To explore the possibility of detecting the oxidative damage lesion Tg, single Tg-containing ODNs (Tg at positions ω13–16) were measured using α-HL pores with a $C_{36}T_4$ control (Figure 2.3). Tg-containing ODNs were more blocking compared to the $C_{36}T_4$ control. The most blocking Tg-containing ODN was Tgω15 (0.9% more blocking than $C_{36}T_4$), which demonstrates position ω15 (from 3' end) is the most sensitive position for Tg detection under +120 mV bias. However, its current blockage levels fall within the range of the native DNA bases (*15*). As for ODNs containing double Tg lesions, the blockage level of $Tg_{\omega 13,16}$ still falls within the range of native bases, while $Tg_{\omega 13,15}$ gave a relatively deeper blockage current level with 1.4% more blocking compared to $C_{36}T_4$, and 0.3% more blocking

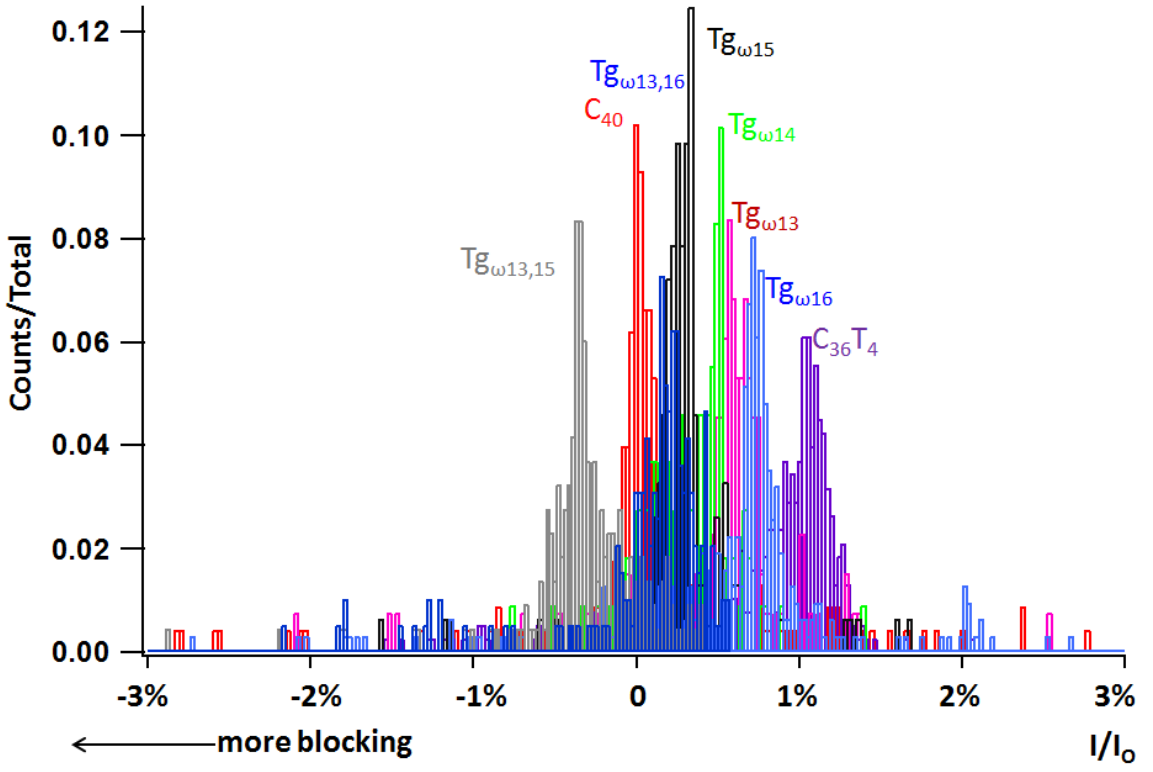


Figure 2.3. Histograms of blockage current percentage (I/I_0) for ODNs containing single or double Tg at position ω13-16. The x-axis is Δ% I/I_0.

compared to C_{40}. This result shows that it is possible to detect tandem lesions through α-HL ion channel if the lesion is close enough and is in the most sensitive sensing position in the β-barrel of the nanopore.

2.3.2 Electrical Signatures of Tg Stereoisomer

The most significant difference between Tg-containing ODNs and T-containing ODNs was observed at position ω15. When Tg was placed at ω15, two populations were observed with a ratio of 85:15 (Figure 2.4). It was reported that solid-phase synthesis of Tg-containing ODNs yields two stereoisomers of *cis*-(5*R*, 6*S*) and *trans*-(5*R*. 6*R*) in a ratio of 87:13 (*5*). This correlates to the population distributions from the nanopore experiment nicely. It has been previously reported that the nanopore can distinguish diastereomers of damaged necleotides including spiroiminodihydantoin (Sp) and guanidinohydantoin (Gh) (*16*).

2.3.3 Electrical Signatures of Bulky Adduct Containing DNA

One other factor that influences the current blockage level is the size, shape, charge, and the hydrophobicity of the molecule (*15,17*). To test the role of hydrophobicity, during ODN preparation, the TBDMS (tert-butyldimethylsilyl) bulky protecting groups were not removed. However, during ODN preparation, TCAA (trichloroacetic acid) was used to deprotect the DMT (dimethoxytrityl) group, which also deprotects the TBDMS group. Thus, three different current blockage levels were generated, corresponding to no TBDMS group (1.0% more blocking than $C_{36}T_4$) and Tg with TBDMS group adducts (2.7% more blocking than $C_{36}T_4$

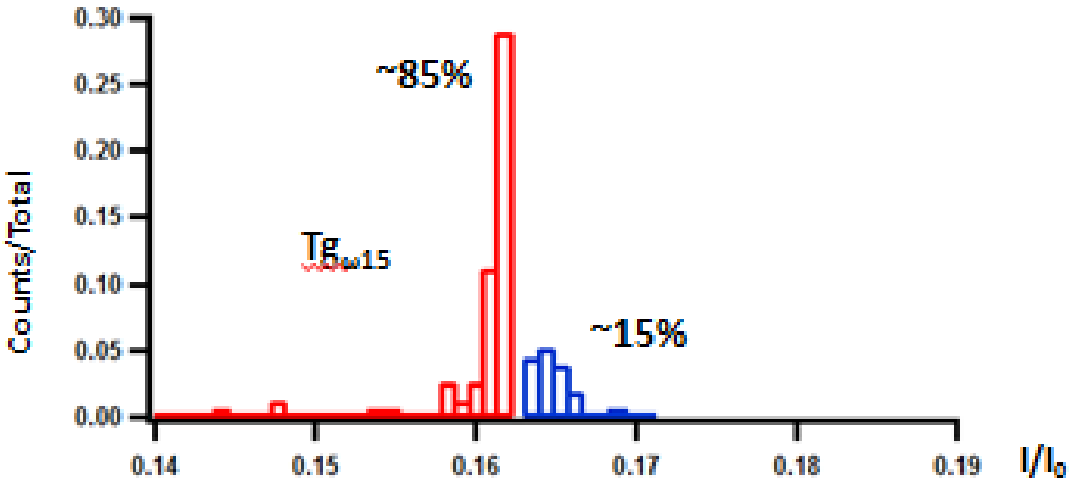


Figure 2.4. Histograms of blockage current percentage (I/I_o) for ODNs containing Tg at position ω15, showing two current populations.

and 3.7% more blocking than $C_{36}T_4$) (Figure 2.5).

2.4 Conclusions

The current blockage levels of ODNs containing a single Tg (position ω13–16) and certain tandem lesions (position ω13,16) fall within the range of native DNA bases. The ODN containing a Tg tandem lesions at positions ω13,15 has a current blockage level outside the level of the native DNA bases, but the difference was small (0.3% more blocking compared to C_{40}). A more sensitive protein nanopore is required to differentiate these current blockage differences during the DNA sequencing process to increase sensitivity and specificity in this method. It was reported that *MspA* has only one sensing zone spanning fewer nucleotides than α-HL, and a larger current level difference between nucleotides was observed (*18*). $Tg_{\omega 15}$ ODN with two-TBDMS groups gives a large blocking level difference and is out of the native-base current blockage range. This result shows that hydrophobic bulky groups contribute to larger current blockages. In the meantime, the bulky group would possibly slow down the ssDNA translocation.

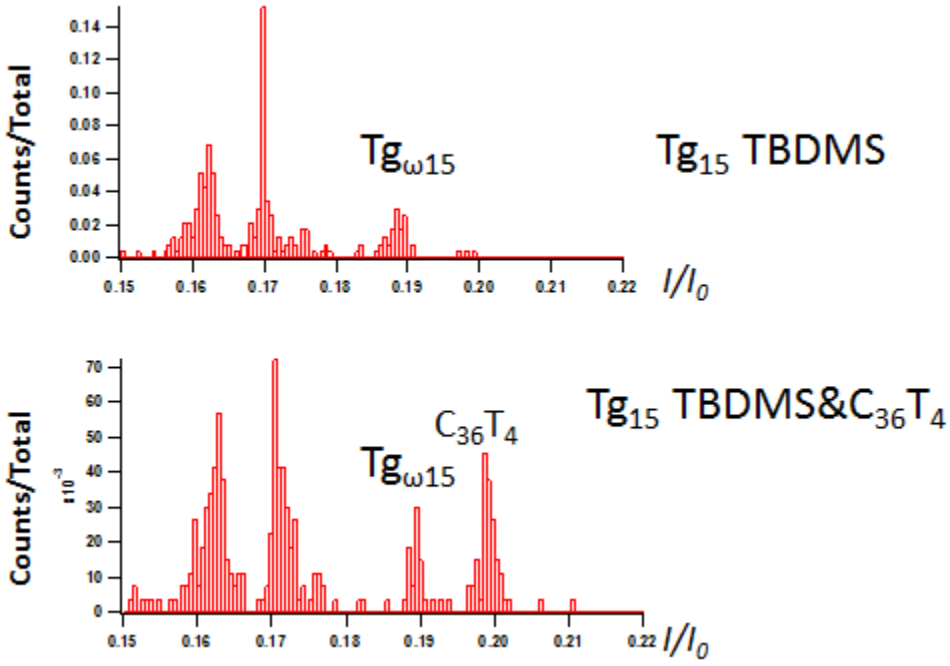


Figure 2.5. Histograms of blockage current percentage (I/I_o) for ODNs containing Tg mono-TBDMS adduct and bis-TBDMS adduct (two left populations) at position ω15.

2.5 References

1. Wallace, S. S. (2002) Biological consequences of free radical-damaged DNA bases, *Free Radic. Biol. Med. 33*, 1–14.

2. Teoule. R., Bonicel, A. (1977) Thymine fragment damage retained in the DNA polynucleotide chain after gamma irradiation in aerated solutions, *Radiat. Res. 72*, 190–200.

3. S. Iida., and Hayatsu., H. (1971) The permanganate oxidation of deoxyribonucleic acid, *Biochim. Biophys. Acta 240*, 370–375.

4. Adelman, R., Saul, R. L., Ames, B. N. (1988) Oxidative damage to DNA: relation to species metabolic rate and life span, *Proc. Natl. Acad. Sci. U.S.A. 85*, 2706–2708.

5. Lustig, M. J., Cadet, J., Boorstein, R. J., and Teebor, G. W. (1992) Synthesis of the diastereomers of thymidine glycol, determination of concentrations and rates of interconversion of their cis-trans epimers at equilibrium and demonstration of differential alkali lability within DNA, *Nucleic Acids Res. 20*, 4839–4845.

6. Kao, J. Y., Goljer, I., Phan, T. A., Bolton, P. H. (1993) Characterization of the effects of a thymine glycol residue on the structure, dynamics, and stability of duplex DNA by NMR, *J. Biol. Chem. 268*, 17787–17793.

7. Ide, H., Kow, Y. W., and Wallace, S. S. (1985) Thymine glycols and urea residues in M13 DNA constitute replicative blocks *in vitro*, *Nucleic Acids Res. 13*, 8035–8051.

8. Breimer, L. H., Lindahl, T. (1984) DNA glycosylase activities for thymine residues damaged by ring saturation, fragmentation, or ring contraction are functions of endonuclease III in *Escherichia coli*, *J. Biol. Chem. 259*, 5543–5548.

9. Jurado, J., Saparbaev, M., Matray, T. J, Greenberg, M. M, Laval, J. (1998) The ring fragmentation product of thymidine C5-hydrate when present in DNA is repaired by the *Escherichia coli* Fpg and Nth proteins, *Biochemistry 37*, 7757–7763.

10. McNulty, J. M., Jerkovic, B., Bolton, P. H., Basu, A. K. (1998) Replication inhibition and miscoding properties of DNA templates containing a site-specific *cis*-thymine glycol or urea residue, *Chem. Res. Toxicol. 11*, 666–673.

11. Azqueta, A., Shaposhnikov, S., and Collins, A. R. (2009) DNA oxidation: Investigating its key role in environmental mutagenesis with the comet assay, *Mutat. Res.-Genetic Toxicol. Environ. Mutagen. 674*, 101–108.

12. Wang, Y. (2002) HPLC isolation and mass spectrometric characterization of two isomers of thymine glycols in oligodeoxynucleotides, *Chem. Res. Toxicol. 15*, 671–676.

13. Stoddart, D., Heron, A. J., Mikhailova, E., Maglia, G., and Bayley, H. (2009) Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore, *Proc. Natl. Acad. Sci. U.S.A. 106*, 7702–7707.

14. Zhang, B., Galusha, J., Shiozawa, P. G., Wang, G., Bergren, A. J., Jones, R. M., White, R. J., Ervin, E. N., Cauley, C. C., and White, H. S. (2007) Bench-top method for fabricating glass-sealed nanodisk electrodes, glass nanopore electrodes, and glass nanopre membranes of controlled size, *Anal. Chem. 79*, 4778–4787.

15. Schibel, A. E., An, N., Jin, Q., Fleming, A. M., Burrows, C. J., White, H. S. (2010) Nanopore detection of 8-oxo-7,8-dihydro-2'-deoxyguanosine in immobilized single-stranded DNA via adduct formation to the DNA damage site, *J. Am. Chem. Soc. 132*, 17992–17995.

16. Wolna, A. H., Fleming, A. M., An, N., He, L., White, H. S., and Burrows, C. J. (2013) Electrical current signatures of DNA base modifications in single molecules immobilized in the alpha hemolysin ion channel, *Isr. J. Chem 53*, 417–430.

17. Mitchell, N., and Howorka, S. (2008) Chemical Tags Facilitate the Sensing of Individual DNA Strands with Nanopores, *Angew. Chem. Int. Ed. 47*, 5565–5568.

18. Butler, T. Z., Pavlenok, M., Derrington, I. M., Niederweis, M., and Gundlach, J. H. (2008) Single-molecule DNA detection with an engineered MspA protein nanopore, *Proc. Natl. Acad. Sci. USA. 105*, 20647–20652.

CHAPTER 3

SINGLE-MOLECULE STUDY OF THE I-MOTIF STRUCTURE THROUGH THE ALPHA-HEMOLYSIN ION CHANNEL

3.1 Introduction

Besides the DNA double helix, DNA oligonucleotides can also adopt many secondary structures that include G-quadruplexes and i-motifs that are stabilized by non-Watson–Crick base pairing (*1, 2*). The i-motif structure results from cytosine-rich stretches and is formed by association of two parallel duplexes in a head to tail orientation through the intercalation of hemiprotonated C^+–C base pairs (Figure 3.1). According to the size of the loop, i-motifs are divided into two classes. The small-loop (Class I) i-motifs have loop sizes of 5'-(2:3/4:2)-3' with four, five, or six C^+–C base pairs. This class includes the VEGF, RET, and Rb sequences; the large-loop (Class II) i-motifs have loop sizes of 5'-(6/8:2/5:6/7)-3', which includes the c-Myc and Bcl-2 sequences (*3*). The formation and stabilization of i-motifs are favored under slightly acidic pH because this maximizes the number of C^+–C base pairs that drives the i-motif formation (N3 of cytosine has a pK_a ~4.6) (*4*). The stability of i-motifs also depends on temperature, salt concentration, and thenucleotides in the loop regions (*5*). Under negative supercoiling, i-motifs can also form under physiological conditions, and in this case, the base interaction

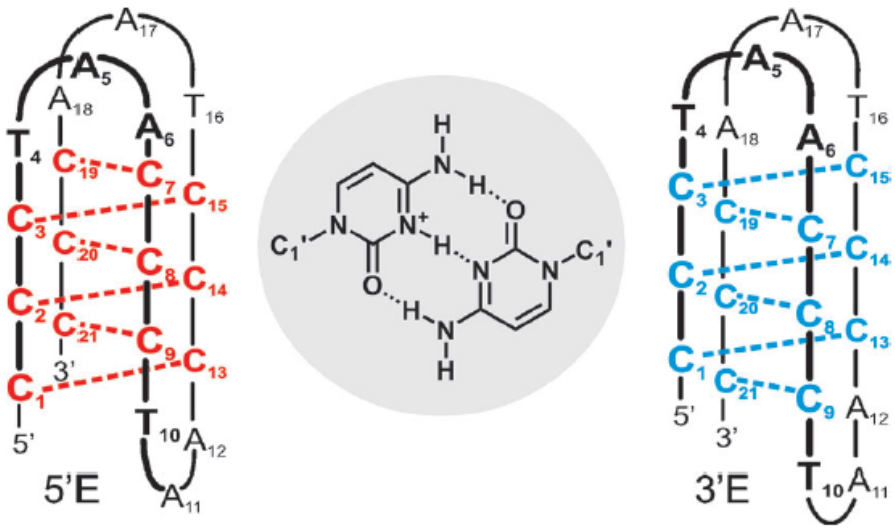


Figure 3.1. Different intercalation topologies of human telomere i-motif DNA. Hemiprotonated C^+–C base pair is shown in the middle of this figure. Figure adapted from Lieblein *et al*.(*9*).

in the loops also drives the formation of a favored i-motif structure (*6*).

There are two cytosine-rich regions in the human genome. The first region is the human telomere, which has a repetitive sequence of 5'-d(CCCTAA)$_n$ (*7*). It was reported that HeLa nuclear protein can specifically bind to the telomeric i-motif sequence, suggesting its role in stabilizing i-motif folds in the human telomere region (*8*). Previously, the Schwalbe group studied the human telomeric sequence 5'-d(CCCTAA)$_3$CCC and found two conformations in equilibrium: 5'E conformation where the closing C^+–C base pair was at 5'-end, and 3'E where the closing C^+-C base pair was at 3'-end (Figure 3.1) (*9*). By changing the nucleotides in the loop regions, the equilibrium ratio between these two conformations change as a result of the difference in T–T base pair and purine nucleobase stacking interactions (*5*).

The other cytosine-rich regions are located in over 50% of human oncogene promoters (*10*). The i-motif sequences in these oncogene promoters have received a lot of attention after many studies have shown their potential biological role, as well as the G-quadruplex structures that form in the complementary strand. Negative supercoiling was demonstrated to facilitate the conversion of duplex DNA to G-quadruplex and i-motif structures in the NHE III1 located in the c-Myc promoter. Because transcription itself can be a source of supercoiling in eukaryotic cells, the formation of G-quadruplex and i-motif structures could regulate the transcription process. This was demonstrated by the fact that once the c-Myc G-quadruplex binding protein, nucleolin, or the drug TMPyP4 has bound these folds, the NM23–H2 protein cannot convert the G-quadruplex and i-motif into their single-stranded forms. Therefore, the transcription process was inhibited. TMPyP4 also

binds i-motif structures, suggesting i-motifs play a similar role as G-quadruplexes in transcriptional regulation (Figure 3.2) (*3, 11, 12*).

Current techniques used to study i-motif conformations include circular dichroism (CD), gel electrophoresis, and nuclear magnetic resonance (NMR) (*9, 11*). The drawback of CD and gel electrophoresis is that these techniques can only provide ensemble information and cannot differentiate different conformations that coexist in the solution. NMR can provide kinetic information of interconversion between different conformations, but it is labor intensive and costly to characterize i-motif structures by NMR.

The alpha-hemolysin (α-HL) ion channel nanopore method has been widely used to study DNA secondary structures such as hairpins and G-quadruplexes (*13*, *14*). Thrombin-binding aptamer (TBA) folds into a two-tetrad G-quadruplex, and the folding and unfolding of this G-quadruplex was detected in a nanopore nanocavity at the single-molecule level (*14*). Kinetics of DNA secondary structure unzipping within α-HL ion channel and protein-DNA interactions can be monitored at the single-molecule level (*15*). Recently, our laboratory showed that the G-quadruplex hybrid 1 and hybrid 2 folds are well differentiated through their interaction with α-HL ion channel (Figure 3.3) (*16*).

The two i-motif sequences in the current study are from the human telomere 5’-d(AACCCTAA)3CCCTA, and from the human RET oncogene promoter region 5’-d(GC CCC CGC CCC GCC CCG CCC CTA). The RET oncogene encodes a receptor-type tyrosine kinase that has been implicated in the development of thyroid cancer (*17*). RET protein levels have also been found to be overexpressed

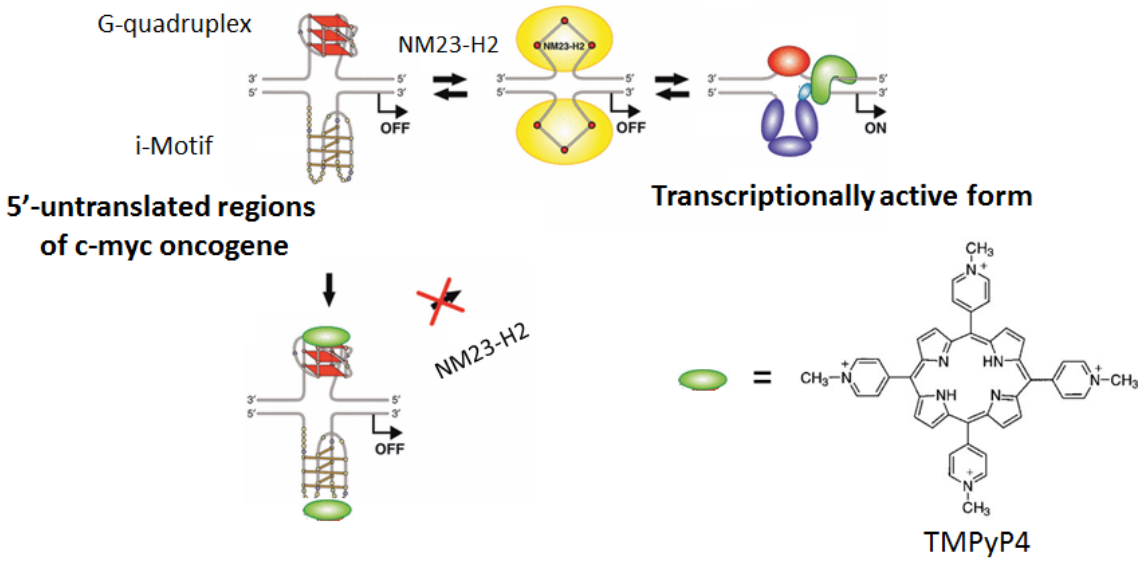


Figure 3.2. A cartoon shows the involvement of NM23–H2 and a G-quadruplex and i-motif interactive compound TMPyP4 in modulating the activation and silencing of the NHE III1 in the c-myc promoter. Figure adapted from Brooks *et al*. (*3*).

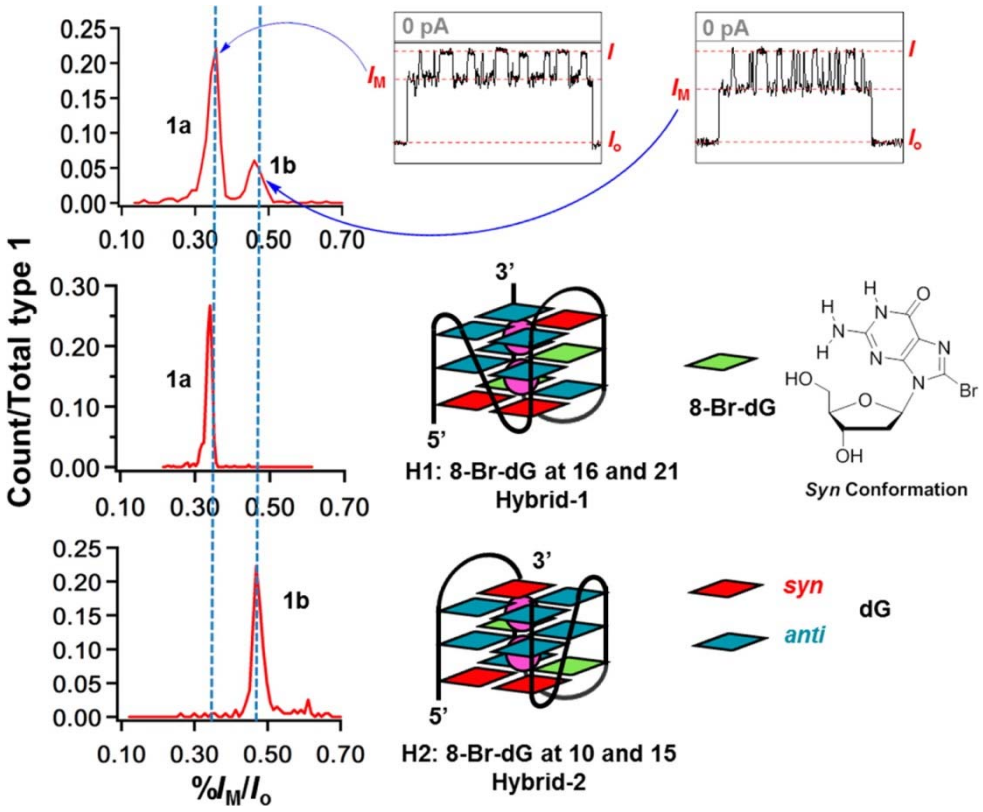


Figure 3.3. Coexistence of hybrid 1 and hybrid 2 folds are resolved using α-HL ion channel. Figure adapted from An *et al*. (*16*).

in medullary thyroid carcinomas and pheochromocytomas (*18*). Most interestingly, the human RET promoter is a TATA-less promoter and has two GC boxes corresponding to three Sp1 binding sites in the proximal promoter region that is essential for basal promoter activity (*19*). Because the i-motif sequence is polymorphic, it can form into many different conformations, in which the α-HL nanopore can provide the population distributions for each fold due to the conformation-dependent current blockage levels. This type of behavior has previously been demonstrated in our laboratory and is discussed below. This will facilitate drug design by other groups because nanopore analysis can provide the most populated i-motif conformation in solution for targeting by small molecules.

3.2 Experimental Section

3.2.1 DNA Preparation, Purification, and Characterization

The telomere i-motif 5'-d(AA CCC TAA CCC TAA CCC TAA CCC TA), RET i-motif 5'-d(GC CCC CGC CCC GCC CCG CCC CTA), and the control 5'-d(GT CTC TGC TCT ATC TCG TCT CTA) oligodeoxynucleotides were synthesized from commercially available phosphoramidites (Glen Research, Sterling, VA) by the DNA-Peptide Core Facility at the University of Utah. Each ODN was cleaved from the synthetic column and deprotected according to the manufacturer's protocols after synthesis. Then the ODNs were purified by an ion-exchange HPLC column. The HPLC method was a linear gradient of 25% to 100% B over 25 minutes while monitoring absorbance at 260 nm (A = 25 mM Tris, 1 M NaCl in 10% CH_3CN/90% ddH_2O, pH 7, B = 10% CH_3CN/90% ddH_2O, flow rate = 1 mL/min). After HPLC

purification, the ODNs were dialyzed against ddH_2O to remove salts and then quantified by UV-Vis.

The i-motif sequences were dissolved in 1 M KCl, 10 mM citric acid-K_2HPO_4 buffer at different pH conditions from 5.0 to 9.2. The solution was heated to 90 ^oC for 15 minutes and then cooled slowly to room temperature. The DNA samples were then stored at -20 ^oC prior to their use.

3.2.2 Chemicals and Materials for Nanopore Measurement

Buffer electrolyte was prepared with 1 M KCl, 10 mM citric acid-K_2HPO_4 buffer solution at different pHs from 5.0 to 9.2. The buffer solution was then filtered with a sterile 0.22 μm Millipore vacuum filter. The phospholipid, 1,2-diphytanoyl-*sn*-glycero-3-phospho-choline (DPhPC), was purchased in a powder form from Avanti Polar Lipids and stored at -20 °C. Before conducting the nanopore experiment, DPhPC was dissolved in decane at 10 mg/mL for use. Wild-type α-hemolysin was dissolved in water at 1 mg/mL and stored at -80 °C. Glass nanopore membranes (GNMs) were fabricated as previously described and used as the support structure for the lipid bilayer (*20*). Before use, GNMs were chemically modified with a 2% (v/v) (3-cyano-propyl) dimethylchlorosilane in acetonitrile to produce a hydrophobic surface.

3.2.3 Current–Time Recordings and Data Analysis

A custom-built, high-impedance, low-noise amplifier and data acquisition system (Electronic Bio Sciences, San Diego, CA) was used for the current-time

recordings. Two Ag/AgCl electrodes were positioned inside and outside of the GNM capillary to apply a voltage. A pressure gauge and a 10 mL gas tight syringe (Hamilton) were attached to the GNM. First the DPhPC solution was painted on the GNM surface, followed by applying a positive pressure to the GNM capillary facilitating reconstitution of a single α-HL ion channel in the lipid bilayer. A good α-HL ion channel was identified by a resistance change from 50–100 GΩ (bilayer resistance) to ~1 GΩ (protein resistance).

After protein reconstitution into the lipid bilayer, the ODN was added to the *cis* compartment of the nanopore cell (5 μM final concentration). For ssDNA translocation experiment, different voltages (+80 mV, +100 mV, and +120 mV) were applied (*trans* vs. *cis*), and the ssDNA translocated through the α-HL pore causing an ion current change from an open channel current (I_0) to a blocking current level (I). A minimum of 800 events were collected for each sample. The current-time traces were filtered at 100 kHz and sampled at 500 kHz. For i-motif capture experiments, since the i-motif is too bulky to enter the constriction zone of α-HL, the blocking current was maintained for 750 ms before -120 mV was applied to eject the ODN from the nanopore for 200 ms. Then the voltage was returned back to +120 mV again to attract the next i-motif (Figure 2.2). A minimum of 300 events were collected for each sample. The current-time traces were filtered at 10 kHz and sampled at 50 kHz.

Events were extracted using QuB (version 1.5.0.31). Histograms of current blockade were plotted using OriginPro (version 8.5.1). Density plots were analyzed with data analysis programs provided by EBS.

3.3 Results and Discussion

3.3.1 Study of Human Telomere i-Motif Structure

3.3.1.1 CD Characterization of the Human Telomere i-Motif Structure

CD spectroscopy is widely used to determine the formation of i-motif structures. The CD spectra of the human telomere i-Motif were collected at different pH values (from 5.3 to 7.9). As shown in Figure 3.4, at acidic pH levels, the spectra have positive peaks at 286 nm and negative peaks at 254 nm. When the pH is greater than 7, the positive peaks decrease and shift to 275 nm. The transition midpoint ~ pH 6.1 was determined by plotting the molar ellipticity at 286 nm versus pH. The experimental result is consistent with a previously published paper (*21*).

3.3.1.2 Electrical Signatures of Human Telomere i-Motif at pH 7.4 and pH 6.3

According to the CD measurements, this i-motif sequence does not fold into an i-motif secondary structure at pH 7.4. The nanopore experiment results showed that the telomere i-motif translocation time was fast, suggesting that it was single-stranded DNA that translocated through the α-HL ion channel (Figure 3.5). At pH 6.3, around 40% of the DNA strands were folded into i-motif structures. According to the experiment result (Figure 3.5), the i-motif event frequency is much lower than it is under pH 5.2. One reason is capture of ssDNA is easier than capture of i-motif due to its size difference. It was reported that the event frequency ratio was around 6:1 between capture of G-quadruplex and triplex structures, and the

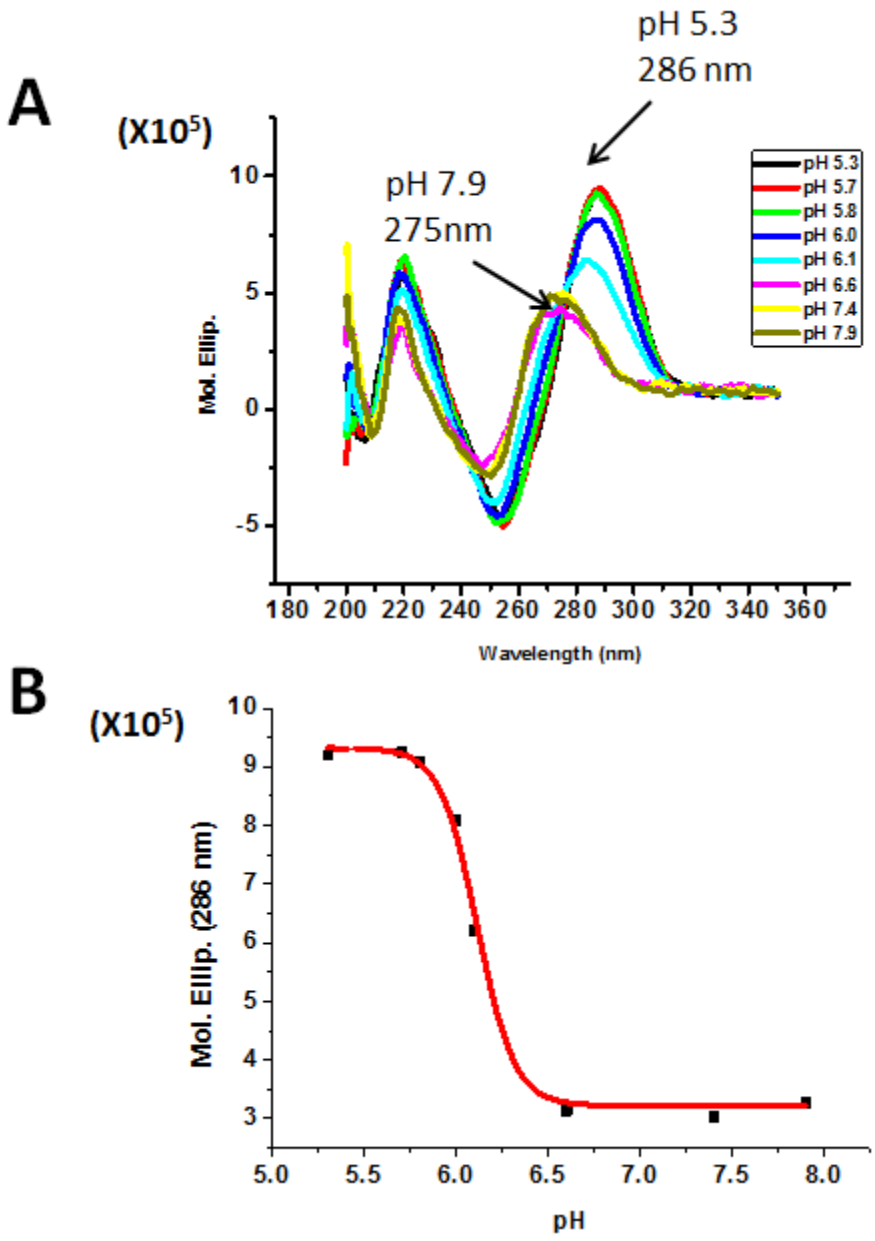


Figure 3.4. CD spectroscopy for human telomere i-motif under different pH (A) and pH dependence of the molar ellipticity at 286 nm (B).

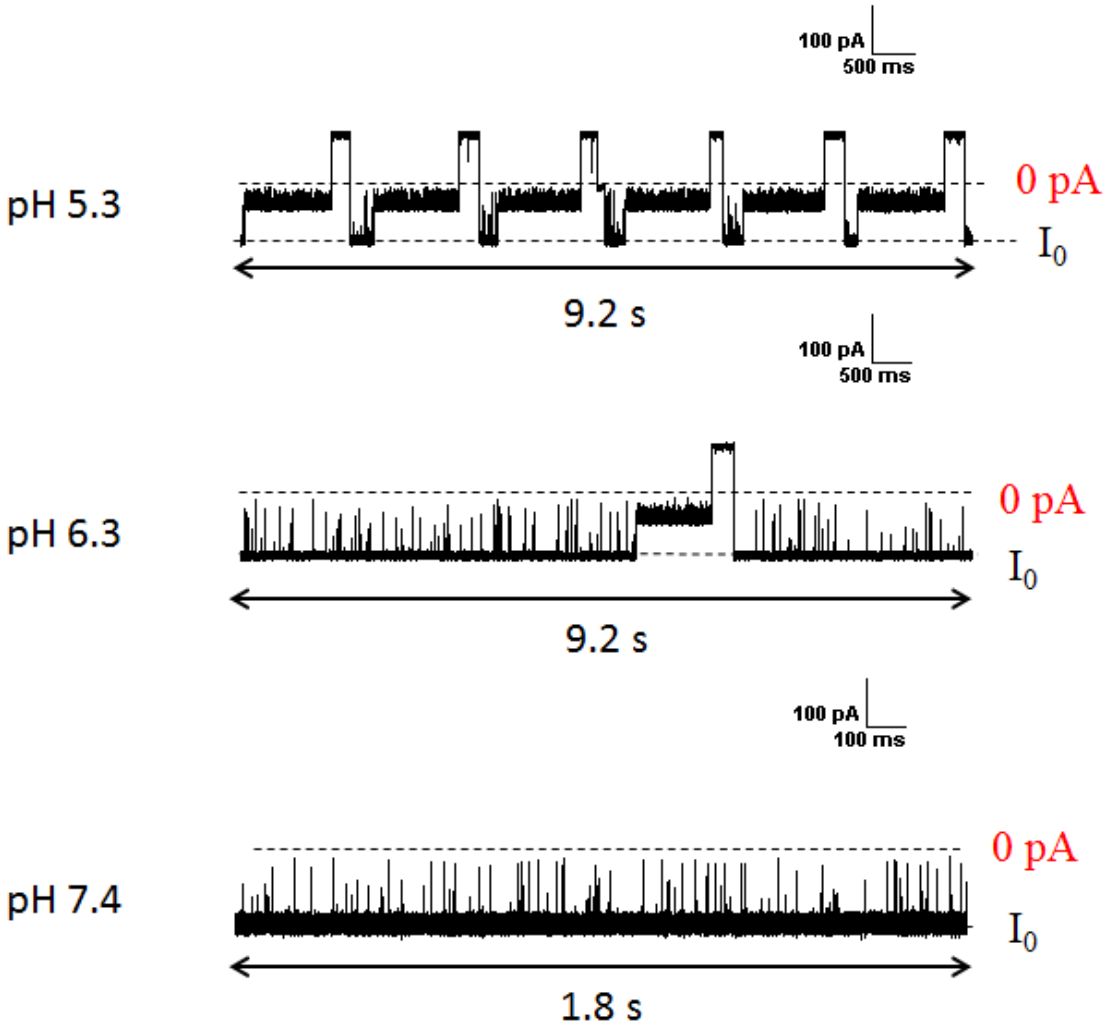


Figure 3.5. Current–time (*i*–*t*) trace of human telomere DNA under pH 5.3, 6.3, and 7.4.

G-quadruplexes are similar in dimensions to the i-motif (*16*).

3.3.1.3 Electrical Signatures of Human Telomere i-Motif at pH 5.3

According to CD measurements, the human telomere sequence folds into an i-motif structure at pH 5.3. As shown in Figure 3.5, the human telomere i-motif in the vestibule of the α-HL ion channel generates long current blockades (most events are longer than 5 seconds, and the DNA was ejected by switching the bias), suggesting the i-motif cannot translocate through the α-HL ion channel. The human telomere i-motif has a dimension of 2.8 nm x 1.6 nm x 1.4 nm (*7*), which makes it fit to the α-HL vestibule, but it cannot translocate through the constriction site (*22*). From the current blockage histogram, as shown in Figure 3.6, there are three current blockage levels: 32% I/I_0, 27% I/I_0 and 8% I/I_0. It is proposed that current blockage at 32% I/I_0 and 27% I/I_0 corresponds to 5'E and 3'E conformation with the ratio ~ 6:1. It was reported that the ratio between 5'E and 3'E conformation is around 4:1 by NMR (*5*). The current blockage around 8% I/I_0 is relatively dispersed, corresponding to 5'-tail or 3'-tail of the i-motif interacting with α-HL constriction zone.

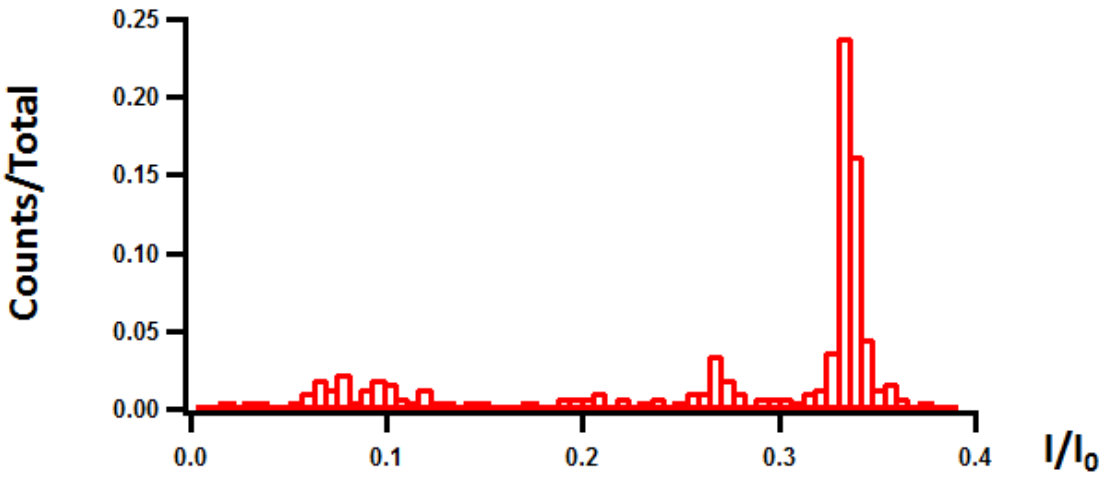


Figure 3.6. Histograms of blockage current percentage (I/I_o) for human telomere i-motif in pH 5.3 buffer.

3.3.2 Study of RET i-Motif Structure

3.3.2.1 CD Characterization of RET i-Motif Structure

The CD spectra of RET were collected at different pH values (from 5.0 to 9.2). As shown in Figure 3.7, at acidic pH levels, the spectra have positive peaks at 288 nm and negative peaks at 264 nm. When the pH is greater than 7, the positive peaks decrease and shift to 282 nm. The transition midpoint near pH 6.2 was determined by plotting the molar ellipticity at 288 nm versus pH. The experimental result is consistent with the published paper by the Hurley laboratory, suggesting the same i-motif sequence used in this project behaves the same as it is in the published paper (*23*).

3.3.2.2 Electrical Signatures of Control DNA Sequence

In order to understand how acidic conditions influence the α-HL ion channel, a control DNA sequence 5'-d(GT CTC TGC TCT ATC TCG TCT CTA) was designed that does not fold into any secondary structure or i-motif at both acidic and neutral conditions, based on M-fold analysis. From the nanopore experimental result as shown in Figure 3.8, at pH 5.2 and pH 5.9, the control DNA sequence translocates through the α-HL ion channel very fast (< 0.1 ms at 120 mV), showing it is single-strand DNA translocation under pH 5.2 and pH 5.9. The control sequence nanopore experiment at pH 7.4 will also be performed later to test if α-HL nanopore behaves similar at acidic condition as in neutral condition.

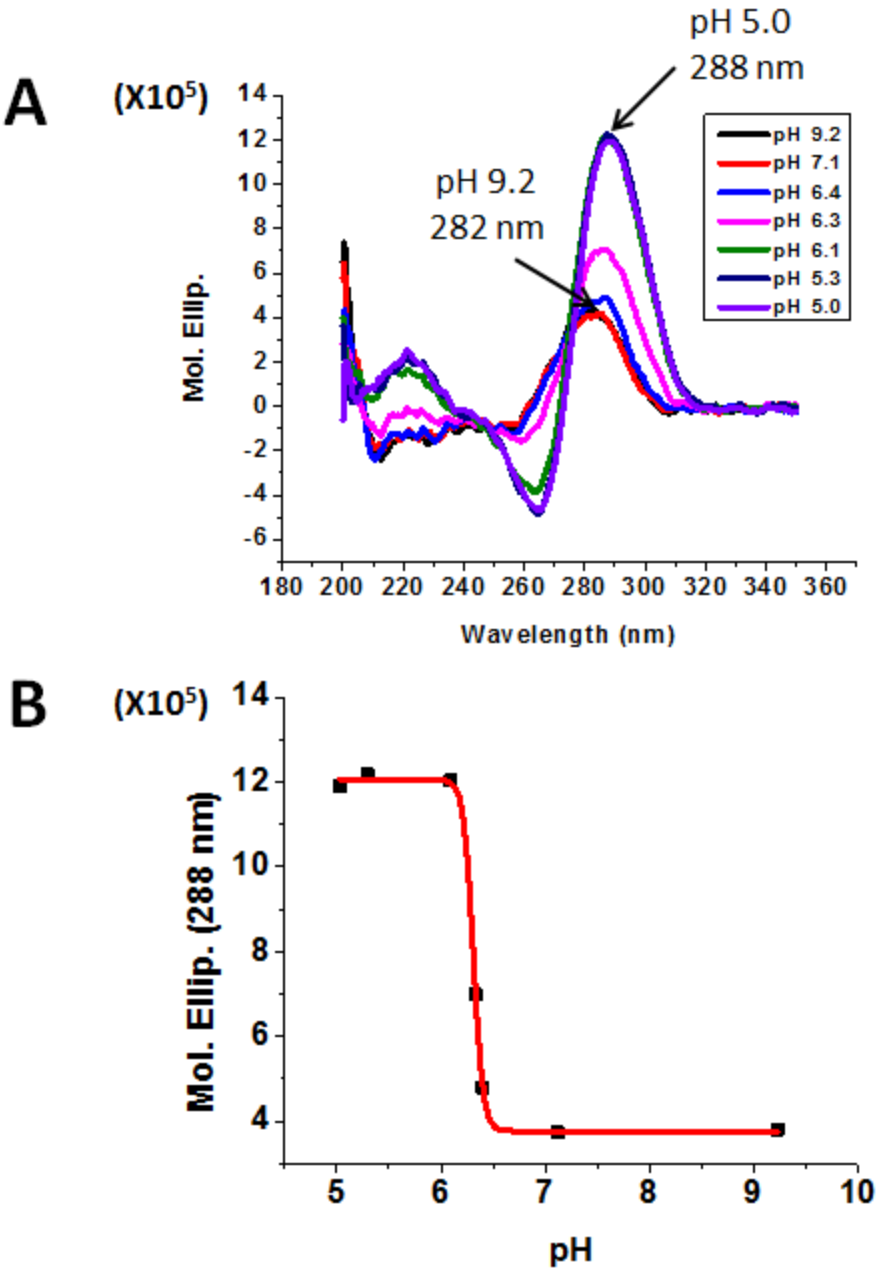


Figure 3.7. CD spectroscopy for RET i-motif under different pH (A) and pH dependence of the molar ellipticity at 288 nm.

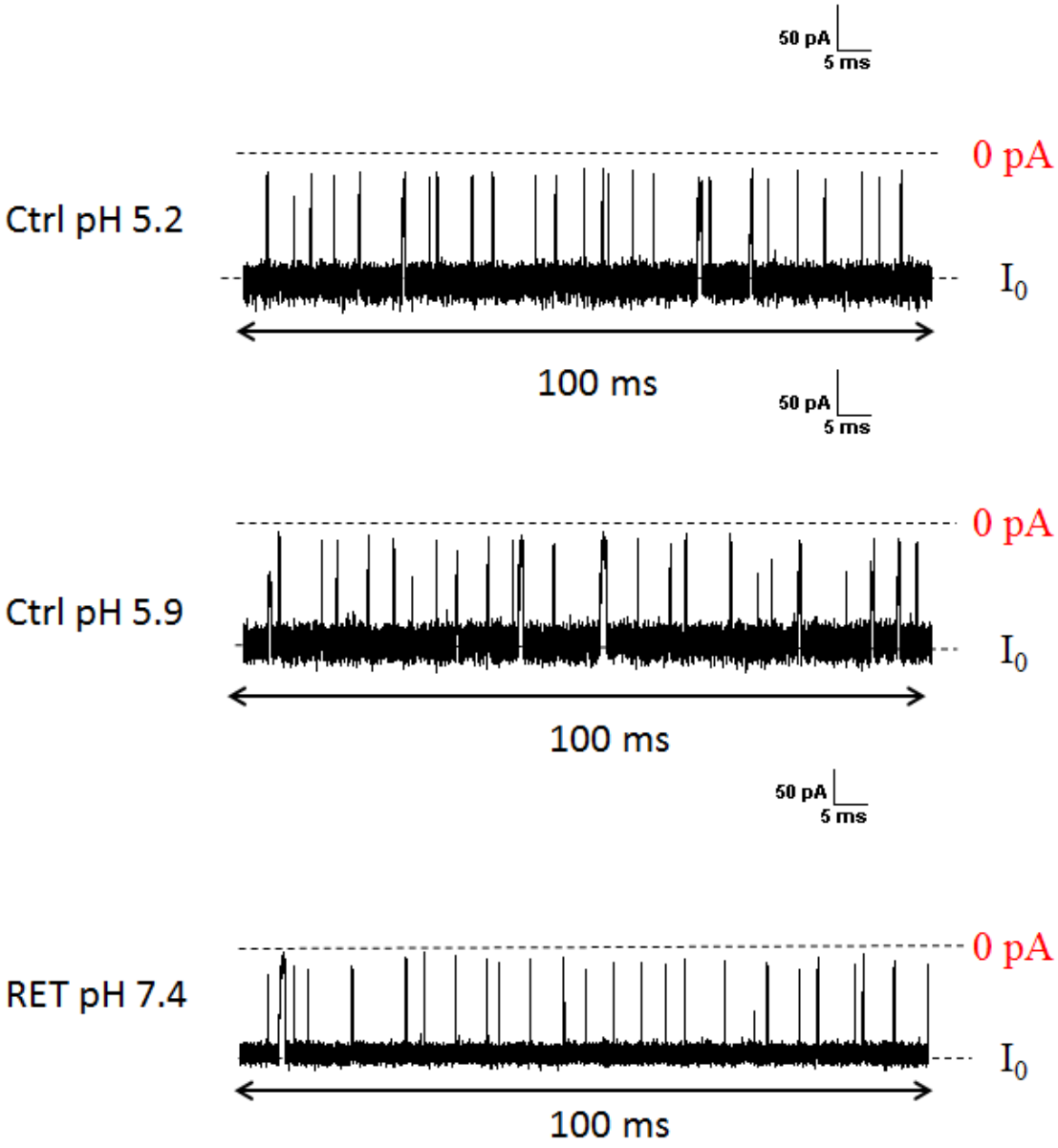


Figure 3.8. Current–time (*i*–*t*) trace of control DNA sequence in pH 5.2, 5.9 buffer and RET sequence in pH 7.4 buffer.

3.3.2.3 Electrical Signatures of RET i-Motif at pH 7.4

According to the CD measurements, this i-motif sequence does not fold into an i-motif secondary structure at pH 7.4. The nanopore experiment results showed that the RET i-motif translocation time was very fast, suggesting that it was single-stranded DNA that translocated through the α-HL ion channel (Figure 3.8).

3.3.2.4 Electrical Signatures of RET i-Motif at Acidic Condition

According to the CD measurements, the RET sequence folds into an i-motif structure at pH 5.2. The nanopore experimental results showed that the i-motif in the vestibule of the α-HL ion channel generates long current blockades (60% of all events are longer than 5 seconds, and DNA was ejected by switching the bias), suggesting the i-motif cannot translocate through the α-HL ion channel. According to the current blockage level, the events are divided into 3 different categories. Type 1 events generate very deep current levels and long blockages (11% I/I_0), this current blockage level is similar to the translocation current blockage level. By comparing the size of the α-HL constriction zone and the i-motif dimensions, it is proposed that one side of the i-motif loop is stuck in the constriction zone and generates a long dwell time and deep current blockage level. Type 2 events generate three middle-current levels and long blockages (27% I/I_0, 31% I/I_0, 36% I/I_0). It is proposed that the i-motif side interacts above the α-HL ion channel constriction zone, and the three different current blockade levels correspond to different i-motif conformations, or different orientations of the i-motif in the vestibule of α-HL ion channel. Type 3 events generate lower-current levels and long

blockages (67% I/I_0). In this event type, there are many deep current level blocking spikes in the middle of the lower current level trace, suggesting that the i-motif sits in the middle of the vestibule of the α-HL ion channel, generating lower-current blockage levels; in the meantime, the i-motif base overhang interacts with the constriction site frequently, generating the deep-current level blocking spikes (Figure 3.9).

3.4 Conclusions

This is the first demonstration of i-motif interaction within the α-HL ion channel. At pH 7.4, the nanopore experiment results showed that the DNA translocation time was very fast, suggesting that it was single-stranded DNA that translocated through the α-HL ion channel. Under acidic pH 5.3, the nanopore experimental results showed the i-motif in the vestibule of the α-HL ion channel generates long current blockades, suggesting the i-motif cannot translocate through the α-HL ion channel. The capture ratio between i-motif form and single-stranded form is shifted under different pH, which is consistent with the CD characterization results. The current blockage histogram can be utilized to analyze different i-motif conformations.

For human telomere i-motif, it was reported that the ratio between 5'E and 3'E conformation is around 4:1 by NMR study (*5*). From nanopore current histogram analysis, two major current blockages at 32% I/I_0 and 27% I/I_0 with a ratio about 6:1 may correspond to 5'E and 3'E conformation.

For RET i-motif, three event types were detected. In type 2 events, the three

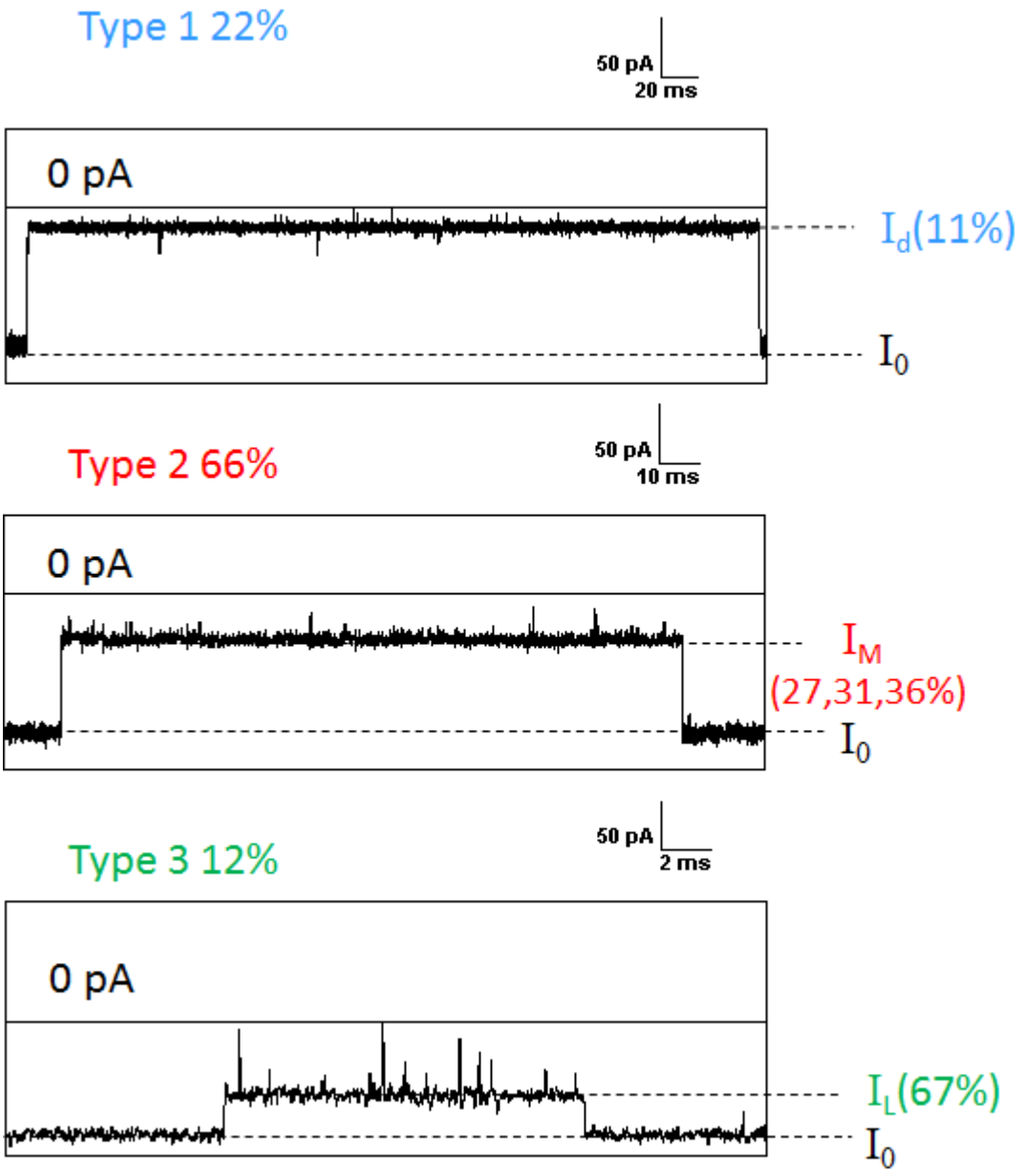


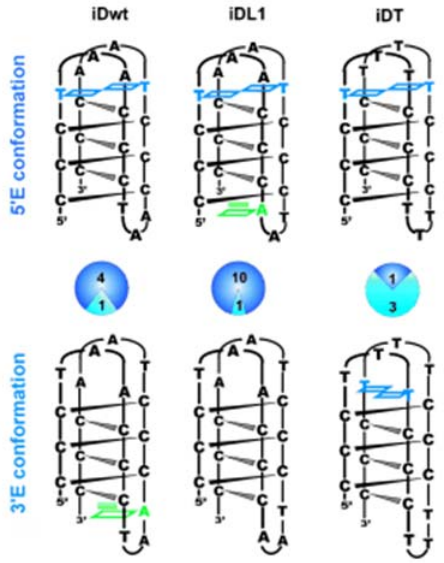
Figure 3.9. Current–time (*i*–*t*) trace of RET i-motif DNA in pH 5.2 buffer.

different current blockade levels correspond to different i-motif conformations, or different orientations of the i-motif in the vestibule of α-HL ion channel. Therefore, a control experiment with a locked conformation of the RET i-motif will be performed. This information will facilitate the drug design, since nanopore analysis can provide the most populated i-motif conformation in solution for targeting by small molecules.

3.5 Future Study

3.5.1 Human Telomere i-Motif 5'E and 3'E Conformations

It was recently reported that changing nucleotides in the loop region of human telomere i-motif shifts the equilibrium of the two conformations: 5'E and 3'E. A wild type i-motif sequence gives a ratio of 4:1 between 5'E and 3'E. By changing the single-loop side nucleotide sequence from TAA to ATT, the equilibrium shifts to 5'E because of the adenosine stacking effect. By changing all three loop sequence from TAA to TTT, the ratio between 5'E and 3'E is reversed to 1:3, as a result of the TT base pair near i-motif core (Figure 3.10) (*5*). The nanopore experiment has shown that a wild type i-motif sequence gives two event populations of 6:1, which is similar to the ratio studies by NMR, presumably an indication of 5'E and 3'E conformations. To confirm this hypothesis, the mutant sequence giving a ratio of 1:3 between 5'E and 3'E by NMR will be used in the nanopore experiment, and we expect a similar population distribution to NMR results on the current blockage histogram.



iDwt 5′-d(AA CCC TAA CCC TAA CCC TAA CCC TA)
iDL1 5′-d(AA CCC TAA CCC AAT CCC TAA CCC TA)
iDT 5′-d(AA CCC TTT CCC TTT CCC TTT CCC TA)

Figure 3.10. Ratios between 5'E and 3'E of i-motif wild type sequence and mutant sequences. Figure adapted from Lieblein *et al*.(*5*).

3.5.2 i-Motif Structure Locking to One Conformation

Denaturing gel electrophoresis gives ensemble information and cannot differentiate different DNA conformations. Although Hurley proposed one RET i-motif conformation based on the gel electrophoresis, there are other possible i-motif conformations that could exist in solution. In fact, by shifting all the bases in the i-motif structure one nucleotide in the 5' to 3' orientation, the same DNA sequence forms a different i-motif conformation. Therefore, based on Hurley's proposed RET i-motif conformation, the cytosines that are not involved in the C^+–C base pairing will be replaced by thymines to enforce the i-motif sequence to form only one conformation. If the nanopore experimental results show only one middle-current level blockade after adding the locked i-motif, it provides support for the proposal that the three different middle-current level blockades correspond to three different i-motif conformations based on the same DNA sequence. The other two middle-current level blockades will also be tested and matched to the other possible i-motif conformations.

3.5.3 i-Motif Stability Study

The time for the i-motif unraveling and translocating through the α-HL ion channel is proportional to the stability of the i-motif structure. In human cells, cytosine spontaneously deaminates to uracil, and this reaction occurs about 500 times per human cell per day (*24*). Since C^+–C base pairing is the base pair unit of the i-motif structure, and the i-motif stability is essential to oncogene transcription repression, it is important to know which cytosine to uracil conversion

has the most effect on the RET i-motif stability. Therefore, each cytosine involved in the C^+–C base pairing will be replaced by uracil, and nanopore experiments will be performed to study the influence of cytosine deamination on unzipping duration with the help of a poly-adenosine tail. These data will be backed up by CD and Tm measurements.

Cytosine is the site of methylation by DNA methyltransferases that yields 5-methylcytosine (mC) (*25*). One role of mC in the upstream of transcriptional start sites is to modulate the expression of genes by inhibiting DNA transcription. These methylation modifications that regulate gene expressions in somatic cells are inheritable and are referred to as epigenetic modifications (*26*). Therefore, it will be very interesting to investigate the effect of base modification on the stability of the i-motif. The cytosine in positions 6, 11, and 16 will be replaced by mC since the CpG sites are vulnerable to hypermethylation. The nanopore experiment will be performed to study their influence on the unzipping duration.

3.5 References

1. Gellert, M., Lipsett, M. N., and Davies, D. R. (1962) Helix formation by guanylic acid, *Proc. Natl. Acad. Sci. U.S.A 48*, 2013–2018.

2. Leroy, J. L., Guéron, M., Mergny, J. L., and Hélène, C. (1994) Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif, *Nucleic Acids Res. 22*, 1600–1606.

3. Brooks, T. A., Kendrick, S., and Hurley, L. (2010) Making sense of G-quadruplex and i-motif functions in oncogene promoters, *FEBS J. 277*, 3459–3469.

4. Jin, K.S., Shin, S. R., Ahn, B., Rho, Y., Kim, S.J., Ree, M. (2009) pH-dependent structures of an i-motif DNA in solution, *J. Phys. Chem. B. 113*, 1852–1856.

5. Lieblein, A. L., Fürtig, B., and Schwalbe, H. (2013) Optimizing the kinetics and thermodynamics of DNA i-motif folding, *ChemBioChem 14*, 1226–1230.

6. Sun, D., Hurley, L. H. (2009) The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: Implications for drug targeting and control of gene expression, *J. Med. Chem. 52*, 2863–2874.

7. Phan, A. T., Guéron, M., and Leroy, J. L. (2000) The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere, *J. Mol. Biol. 299*, 123–144.

8. Marsich, E., Piccini, A., Xodo, L. E., and Manzini, G. (1996) Evidence for a HeLa nuclear protein that binds specifically to the single-stranded d(CCCTAA)$_n$ telomeric motif, *Nucleic Acids Res. 24*, 4029–4033.

9. Lieblein, A. L., Buck, J., Schlepckow, K., Fürtig, B., and Schwalbe, H. (2012) Time-resolved NMR spectroscopic studies of DNA i-motif folding reveal kinetic partitioning, *Angew. Chem. Int. Ed. 51*, 250–253.

10. Huppert, J. L., and Balasubramanian, S. (2006) G-quadruplexes in promoters throughout the human genome, *Nucleic Acids Res. 35*, 406–413.

11. Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription, *Proc. Natl. Acad. Sci. U.S.A 99*, 11593–11598.

12. Fedoroff, O. Y., Rangan, A., Chemeris, V. V., and Hurley, L. H. (2000) Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism, *Biochemistry 39*, 15083–15090.

13. DeGuzman, V. S., Lee, C. C., Deamer, D. W., and Vercoutere, W. A. (2006) Sequence-dependent gating of an ion channel by DNA hairpin molecules, *Nucleic Acids Res. 34*, 6425–6437.

14. Shim, J. W., Tan, Q., and Gu, L. (2009) Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity, *Nucleic Acids Res. 37*, 972–982.

15. Mathé, J., Visram, H., Viasnoff, V., Rabin, Y., and Meller, A. (2004) Nanopore Unzipping of Individual DNA Hairpin Molecules, *Biophys. J. 87*, 3205–3212.

16. An, N., Fleming, A. M., and Burrows, C. J. (2013) Interactions of the Human telomere sequence with the nanocavity of the α-hemolysin ion channel reveal structure-dependent electrical signatures for hybrid folds, *J. Am. Chem. Soc. 135*, 8562–8570.

17. Kodama, Y., Asai, N., Kawai, K., Jijiwa, M., Murakumo, Y., Ichihara, M., and Takahashi, M. (2005) The RET proto-oncogene: A molecular therapeutic target in thyroid cancer, *Cancer Sci. 96*, 143–148.

18. Veit, C. (2004) Activation of phosphatidylinositol 3-kinase and extracellular signal-regulated kinase is required for glial cell line-derived neurotrophic factor-induced migration and invasion of pancreatic carcinoma cells, *Cancer Res. 64*, 5291–5300.

19. Andrew S.D., Delhanty, P. J., Mulligan, L.M., Robinson, B.G. (2000) Sp1 and Sp3 transactivate the RET proto-oncogene promoter, *Gene. 256*, 283–291.

20. Zhang, B., Galusha, J., Shiozawa, P. G., Wang, G., Bergren, A. J., Jones, R. M., White, R. J., Ervin, E. N., Cauley, C. C., and White, H. S. (2007) Bench-top method for fabricating glass-sealed nanodisk electrodes, glass nanopore electrodes, and glass nanopore membranes of controlled size, *Anal. Chem. 79*, 4778–4787.

21. Kaushik, M., Suehl, N., and Marky, L. A. (2007) Calorimetric unfolding of the bimolecular and i-motif complexes of the human telomere complementary strand, d(C3TA2)4, *Biophys. Chem. 126*, 154–164.

22. Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., Gouaux, J. E. (1996) Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore, *Science. 274*, 1859–1866.

23. Guo, K., Pourpak. A., Beetz-Rogers, K., Gokhale, V., Sun, D., Hurley, L.H. (2007) Formation of pseudosymmetrical G-quadruplex and i-motif structures in the proximal promoter region of the RET oncogene, *J. Am. Chem. Soc.* 129, 10220–10228.

24. Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature 362*, 709–715.

25. Santi, D. V., Garrett, C. E, Barr, P. J. (1983) On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs, *Cell. 33*, 9–10.

26. Klose, R. J., and Bird, A. P. (2006) Genomic DNA methylation: The mark and its mediators, *Trends Biochem. Sci. 31*, 89–97.